

O200
OIPE

RAW SEQUENCE LISTING

DATE: 07/05/2000

PATENT APPLICATION: US/09/599,087

TIME: 18:53:55

Input Set : A:\00450seq.txt

Output Set: N:\CRF3\07052000\I599087.raw

3 <110> APPLICANT: Polverino, Anthony J.
 4 Luethy, Roland
 6 <120> TITLE OF INVENTION: Secreted Epithelial Colon Stromal-1 Molecules and Uses
 7 Thereof
 9 <130> FILE REFERENCE: 00-450
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/599,087
 C--> 12 <141> CURRENT FILING DATE: 2000-06-21
 14 <160> NUMBER OF SEQ ID NOS: 21
 16 <170> SOFTWARE: PatentIn Ver. 2.0
 18 <210> SEQ ID NO: 1
 19 <211> LENGTH: 744
 20 <212> TYPE: DNA
 21 <213> ORGANISM: Mus musculus
 23 <220> FEATURE:
 24 <221> NAME/KEY: CDS
 25 <222> LOCATION: (38)..(274)
 27 <400> SEQUENCE: 1
 28 gcttcctccc taggcgtgag actccggctc cttcact atg aga ctt cta gcc ctt 55
 29 Met Arg Leu Leu Ala Leu
 30 1 5
 32 tcc ggt ctg ctc tgc atg ctg ctc ctc tgt ttc tgc att ttc tcc tca 103
 33 Ser Gly Leu Leu Cys Met Leu Leu Leu Cys Phe Cys Ile Phe Ser Ser
 34 10 15 20
 36 gaa ggg aga aga cat cct gcc aag tcc ttg aaa ctc agg cgc tgc tgt 151
 37 Glu Gly Arg Arg His Pro Ala Lys Ser Leu Lys Leu Arg Arg Cys Cys
 38 25 30 35
 40 cac cta tct cct aga tcc aag ctg aca acc tgg aaa gga aac cac aca 199
 41 His Leu Ser Pro Arg Ser Lys Leu Thr Thr Trp Lys Gly Asn His Thr
 42 40 45 50
 44 agg ccc tgc aga ctc tgc aga aac aag cta cca gtc aag tca tgg gtg 247
 45 Arg Pro Cys Arg Leu Cys Arg Asn Lys Leu Pro Val Lys Ser Trp Val
 46 55 60 65 70
 48 gtg cct ggg gct ctc cca cag ata tag ggctcctcc gccagatga 294
 49 Val Pro Gly Ala Leu Pro Gln Ile
 50 75
 52 agcgttgatg ccagatgtg gagacaccag aagcatacac actatgttgc cttgcccctt 354
 54 gccaatgagc tgtgacctg gaatgcttca cttcagacat cagggcggat ggattgcaga 414
 56 attccaagtc ctcatccaa aggtgtcacc aaccttcaga gtcactaagg tccaggctca 474
 58 gcccaagat caccatggct cctccagagt aaaagtccaa gattccaact gtgggagcta 534
 60 cagatccaga gactttcaag ctgactagag tgcagagaag caagacctca gtgtgatcag 594
 62 ccgagactac agcatcttgg gaacctcag tcagcccca acccctaaca ctttaaccact 654
 64 ggtctccaaa ccaacacctg taacttctca atgaaatcat caggaggata ccaaaagaaa 714
 66 taaaccataa atcagcatac acactaaaaa 744
 69 <210> SEQ ID NO: 2
 70 <211> LENGTH: 78
 71 <212> TYPE: PRT
 72 <213> ORGANISM: Mus musculus

ENTERED

RAW SEQUENCE LISTING DATE: 07/05/2000
 PATENT APPLICATION: US/09/599,087 TIME: 18:53:55

Input Set : A:\00450seq.txt
 Output Set: N:\CRF3\07052000\I599087.raw

```

74 <400> SEQUENCE: 2
75 Met Arg Leu Leu Ala Leu Ser Gly Leu Leu Cys Met Leu Leu Leu Cys
76   1           5           10           15
78 Phe Cys Ile Phe Ser Ser Glu Gly Arg Arg His Pro Ala Lys Ser Leu
79           20           25           30
81 Lys Leu Arg Arg Cys Cys His Leu Ser Pro Arg Ser Lys Leu Thr Thr
82           35           40           45
84 Trp Lys Gly Asn His Thr Arg Pro Cys Arg Leu Cys Arg Asn Lys Leu
85           50           55           60
87 Pro Val Lys Ser Trp Val Val Pro Gly Ala Leu Pro Gln Ile
88   65           70           75
91 <210> SEQ ID NO: 3
92 <211> LENGTH: 54
93 <212> TYPE: PRT
94 <213> ORGANISM: Mus musculus
96 <400> SEQUENCE: 3
97 Arg Arg His Pro Ala Lys Ser Leu Lys Leu Arg Arg Cys Cys His Leu
98   1           5           10           15
100 Ser Pro Arg Ser Lys Leu Thr Thr Trp Lys Gly Asn His Thr Arg Pro
101           20           25           30
103 Cys Arg Leu Cys Arg Asn Lys Leu Pro Val Lys Ser Trp Val Val Pro
104           35           40           45
106 Gly Ala Leu Pro Gln Ile
107   50
110 <210> SEQ ID NO: 4
111 <211> LENGTH: 806
112 <212> TYPE: DNA
113 <213> ORGANISM: Homo sapiens
115 <220> FEATURE:
116 <221> NAME/KEY: CDS
117 <222> LOCATION: (29)..(274)
119 <400> SEQUENCE: 4
120 ggaacgaggg aaaatctgcc ttctcacc atg agg ctt cta gtc ctt tcc agc   52
121           Met Arg Leu Leu Val Leu Ser Ser
122           1           5
124 ctg ctc tgt atc ctg ctt ctc tgc ttc tcc atc ttc tcc aca gaa ggg   100
125 Leu Leu Cys Ile Leu Leu Cys Phe Ser Ile Phe Ser Thr Glu Gly
126   10           15           20
128 aag agg cgt cct gcc aag gcc tgg tca ggc agg aga acc agg ctc tgc   148
129 Lys Arg Arg Pro Ala Lys Ala Trp Ser Gly Arg Arg Thr Arg Leu Cys
130   25           30           35           40
132 tgc cac cga gtc cct agc ccc aac tca aca aac ctg aaa gga cat cat   196
133 Cys His Arg Val Pro Ser Pro Asn Ser Thr Asn Leu Lys Gly His His
134           45           50           55
136 gtg agg ctc tgt aaa cca tgc aag ctt gag cca gag ccc cgc ctt tgg   244
137 Val Arg Leu Cys Lys Pro Cys Lys Leu Glu Pro Glu Pro Arg Leu Trp
138           60           65           70
140 gtg gtg cct ggg gca ctc cca cag gtg tag cactcccaaa gcaagactcc   294
141 Val Val Pro Gly Ala Leu Pro Gln Val

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/599,087
 DATE: 07/05/2000
 TIME: 18:53:55

Input Set : A:\00450seq.txt
 Output Set: N:\CRF3\07052000\I599087.raw

```

142          75          80
144 agacagcgga gaacctcatg cctggcacct gaggtacceca gcagcctcct gtctcccott 354
146 tcagccttca cagcagtgag ctgcaatggt ggagggcttc atctcgggct gcaaggaccc 414
148 tgggaaagt ccagaactcc acgtccttgt ctcaattgtg ccatcaactt tcagagctat 474
150 catgagccaa cctcacccca cagggcctca gtcgccacca tgtgggcctc tccagtgcaa 534
152 accaccgagc attccaccat gaccgggtcac agctacaaat ccagagacca tcaatcctgc 594
154 tagagtgcag ggtggcaagc acccaagggt ggctgaccaa gactgcagag tctcctccat 654
156 cttcaggtcc attcagcctc ctggcattta actaccagca tccagtggtc cccaaggaat 714
158 cccttcctag cctcctgaca ttagtctgct ggaaagagca tccaaacaaa caagtaataa 774
160 ataaataaat aaactcaatg cagacacaaa aa                                806
163 <210> SEQ ID NO: 5
164 <211> LENGTH: 81
165 <212> TYPE: PRT
166 <213> ORGANISM: Homo sapiens
168 <400> SEQUENCE: 5
169 Met Arg Leu Leu Val Leu Ser Ser Leu Leu Cys Ile Leu Leu Leu Cys
170 1          5          10          15
172 Phe Ser Ile Phe Ser Thr Glu Gly Lys Arg Arg Pro Ala Lys Ala Trp
173 20          25          30
175 Ser Gly Arg Arg Thr Arg Leu Cys Cys His Arg Val Pro Ser Pro Asn
176 35          40          45
178 Ser Thr Asn Leu Lys Gly His His Val Arg Leu Cys Lys Pro Cys Lys
179 50          55          60
181 Leu Glu Pro Glu Pro Arg Leu Trp Val Val Pro Gly Ala Leu Pro Gln
182 65          70          75          80
184 Val
187 <210> SEQ ID NO: 6
188 <211> LENGTH: 57
189 <212> TYPE: PRT
190 <213> ORGANISM: Homo sapiens
192 <400> SEQUENCE: 6
193 Lys Arg Arg Pro Ala Lys Ala Trp Ser Gly Arg Arg Thr Arg Leu Cys
194 1          5          10          15
196 Cys His Arg Val Pro Ser Pro Asn Ser Thr Asn Leu Lys Gly His His
197 20          25          30
199 Val Arg Leu Cys Lys Pro Cys Lys Leu Glu Pro Glu Pro Arg Leu Trp
200 35          40          45
202 Val Val Pro Gly Ala Leu Pro Gln Val
203 50          55
206 <210> SEQ ID NO: 7
207 <211> LENGTH: 77
208 <212> TYPE: PRT
209 <213> ORGANISM: Rattus norvegicus
211 <400> SEQUENCE: 7
212 Met Arg Leu Leu Thr Leu Ser Gly Leu Phe Phe Met Leu Phe Leu Cys
213 1          5          10          15
215 Leu Cys Val Leu Ser Ser Glu Gly Arg Lys Arg Pro Ala Lys Phe Pro
216 20          25          30
218 Lys Leu Arg Pro Arg Cys His Leu Ser Pro Arg Ser Lys Pro Ile Thr

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/599,087
 DATE: 07/05/2000
 TIME: 18:53:55

Input Set : A:\00450seq.txt
 Output Set: N:\CRF3\07052000\I599087.raw

```

219          35          40          45
221 Trp Lys Gly Asn His Thr Arg Pro Cys Arg Pro Cys Arg Lys Leu Glu
222          50          55          60
224 Ser Asn Ser Trp Val Val Pro Gly Ala Leu Pro Gln Ile
225 65          70          75
228 <210> SEQ ID NO: 8
229 <211> LENGTH: 4159
230 <212> TYPE: DNA
231 <213> ORGANISM: Homo sapiens
233 <220> FEATURE:
234 <221> NAME/KEY: unsure
235 <222> LOCATION: (160)..(169)
237 <220> FEATURE:
238 <221> NAME/KEY: unsure
239 <222> LOCATION: (3884)..(3893)
241 <220> FEATURE:
242 <221> NAME/KEY: exon
243 <222> LOCATION: (1)..(69)
245 <220> FEATURE:
246 <221> NAME/KEY: exon
247 <222> LOCATION: (2627)..(2725)
249 <220> FEATURE:
250 <221> NAME/KEY: exon
251 <222> LOCATION: (4079)..(4159)
253 <400> SEQUENCE: 8
254 atg agg ctt cta gtc ctt tcc agc ctg ctc tgt atc ctg ctt ctc tgc 48
256 ttc tcc atc ttc tcc aca gaa ggtaggcag ccccccagggt gcagatccct 99
258 gagcaggatt tcagcatctg ggaagactct gatcaggatt tgttgagggg caggccttgg 159
W--> 260 nnnnnnnnnn cgcgcggtact tccagccccg tgggtgaagac gaaagagggc tctttctcct 219
262 gaacctatag gtttggggct caggactgcc tgcagggtggc ttggggggttc cattcacagc 279
264 cctctgcacc ccaaatacat acccagccta agtaaatggtg tgtgttcgcc atgcaaacac 339
266 acatacaacc tctcagctag attactgtgc ttaagtcccta cctatctaga atttctggag 399
268 ccattctctt gtacttgggt catgcttggg acagagtaaa ttagtggttg gcaaatgaat 459
270 acattaatta gtagaccatc taagtctgaa catcccaaaa cctcatgccc agaaaaatc 519
272 catgagcagc tgaaatgaag gtgtgtgtgg tagggagggtg ggggtatgtt atgcatgttt 579
274 agaaggggac accatctttt tacctctata gatatgaata ttagctcttc ttgccctttt 639
276 ttcttttttc tttttttttt tttttttgag atggagtctt gctctgtcac ccaggctgga 699
278 gtgcagtggc gctatctcag ctcaactgcaa tctccgcctc ctgggttcaa gcaattctct 759
280 gcctcagcct cccaagttagc tgagattaca ggtgcccacc accaagccca gctaattttt 819
282 gtatttttag tacagacagg ttccaccatc ttggccaggc tggctctgaa ctccctaacct 879
284 cgtaatcctc ccacctcggc ctcccaagt gctgggatta caggcgtgag ccaccatgcc 939
286 tggctgcctt tcttgattca gatagctgag tgtttcaatc catttttctc ttgtctaac 999
288 ctctagaaac tgctacatt tattttttgt ttagtggtt atggttactc aaacttttgg 1059
290 gtggggggag ctggagctat agaaatatat aaagagaaga aaaacactca attccatgat 1119
292 tcaagagtag ccatgttcaa cattttgttt atttccttgc atgtagaatt tttaaaaatt 1179
294 aattgatgta cctatatgtt caaggttata tcttttttat ttatcactat atattattgt 1239
296 ataatacccc aaaatgctta tgattgaaga tattctggaa gcatttaciaa cccagtgtca 1299
298 gcagcagcca tctctgagta gtgggattat aacaagtgtt tgttttaciaa agtttctgag 1359
300 atgaaaatgt cccacatata taataaggaa aacagtgtat agaattcctc ataaacacag 1419

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/599,087
 DATE: 07/05/2000
 TIME: 18:53:55

Input Set : A:\00450seq.txt
 Output Set: N:\CRF3\07052000\I599087.raw

```

302 cccgtgacat gcaatttate agacctctat ttttggacat gttggagggt gccagtgata 1479
304 ccctagtgac aattaaatga ggatagatac cttcccccat aaagtcttct atccatttag 1539
306 gactatctgt agcaaaactct tgaagtagca ttaatcaact aatattttca ggtataactt 1599
308 gctacaagtg aacgtactat gatgaattta catgcttaga catttagata gttcacaatt 1659
310 gtgtgctttt ctttttttga agcaagatct tgctctcttg cccagggtcg agtgagtggt 1719
312 catgaccacg gctcagtgca ggcttgactt ccagggtcga agcaatactc gcacctcagg 1779
314 ttttccagta gctgggaaaa cagggtgcga ccacaatgcc ctgctaattt ttaaaatttt 1839
316 ttgcagagac gaggtctctc taagtgtccc aggtctgtct tgaacttctg gactcaagcc 1899
318 atoctccac cttggcctcc cagagtctca ggatcacagg catgagccac cacacctggc 1959
320 ctacttttga cattttaatt atgtggtaaa aggtatata gtacataaag tatgtccttt 2019
322 attcaggctt tttttctttt tttctttttt ttattttttt gagacgaagt ttttgccttt 2079
324 gttgtccagg ctggagtgtg atggcatgct cttggctcac cacaacctcc gctccccggg 2139
326 ttcaagtgtg tctcctgcct caacctcctg agtagctggg attacaggca tgcaccaaca 2199
328 tgccaggctg attttgtatt tttagtagag atgggtgttc tccatgttgg tcaggctggg 2259
330 ctgcaaacat cgacctcaag tgatccgccc acctcagcct cccaaagagc taggattaca 2319
332 ggcagagacc accacaccca gctcagggtt tattttctta ggctagattg ccaaggggag 2379
334 aattatttat tcaaaagaac tactttattg acaggaatct gaaaaggatg tgttttggg 2439
336 ccattgtgtt cccaacattg ttatttctga aaagtaaata acaacaaggc ccactctttc 2499
338 cctaggacct ctgtagcctt ggctcctcct gagtttctct ggataaatat tcttgagccc 2559
340 tgtgccttgg aaggggaagc tcaactcacag acaagcccac taaagacagt ctctcttctt 2619
342 ttgtgtc cac cct cag gga aga ggc gtc ctg cca agg cct ggt cag gca 2668
344 gga gaa cca ggc tct gct gcc acc gag tcc cta gcc cca act caa caa 2716
346 acc tga aag gtaagtaccc ccacctcgtc cagactgtgg ggcagaagtt 2765
348 ctacagtggc catgggacca gccacacaca ctgatcagcc cccaccatg gctggcatca 2825
350 ggctctggct gggaggacat cttgttttg ttgattaatt tgttgactcc ccccaaaaag 2885
352 tcaacaaatt aatcatttta aactgaatac attctgccat ggaaaaaaag caggatgcaa 2945
354 tttagcagatg ttgtgtgaa acacacttac tttaggtgga aggtgtctga gcaggtgaca 3005
356 tttatgagac ctggctcatt tatgagccag gagcctggct gagcctgtg gaggtggggc 3065
358 atgcaggcag aggaggcagc aagggtgaag ggcaagagtg gggtatggaa gacagatggg 3125
360 agcagggtct gagaggtact ccagaaagct aaggacaaa gctgcctgtg aacctgtgg 3185
362 acctggggca cagatcagca tgcaggtcac cagcagggga gtgggctga ggtccagag 3245
364 agccatagct tggcaggaga taaggcagcc ccagagatgc cagcaggcag catccaggct 3305
366 gcatgaccag aacgagggcc agaagagcaa ggctgcctc tccctgaggc ctggggacac 3365
368 tgggaggcct gtggcgga ggcaccaagc caggagggtc gcgggcaccc agttccctgc 3425
370 acaggggctg caggcccgaga gcagatatct actggagtgg cccagcccag gtggaagggt 3485
372 caggctgctg gagcttgggt agggcaggca gatcccaaag gggagactgt ggacctgag 3545
374 tcagacagcc tgacaccaac ctggggctcc tgctgaaact ctgcagcccc agtgccact 3605
376 ctcaagagge tgaggaggtc ccggccccac ttgctcctct gcggccatgg cccatgggg 3665
378 ccattgaccag cggcgagacc tccatgcctt tccagctac caaggggatg ctgagctgtg 3725
380 atgcaggaga gggatagagg gaggaagcaa gacagcatga ctccagccgc agaccttctc 3785
382 ccggagatgc tgacagccct ttcttccaaa ctggcatcac acccagccgg ccaggataaa 3845
W--> 384 aataaccagc tcgtcttccac cagggtctga aggatccnn nnnnnnnnca cgaagagccc 3905
386 cttctgggct tccagggaaa agcataagat ctaattcttg ctttgaaatt tttttttaa 3965
388 tgtgtttgaa aatgcaactt aattgtgttt tctctctct cccacaacc ttgctctgac 4025
390 ctgcccattt tctgtcctt gtccctcttg tctactcatt gctcctccca gga cat 4081
392 cat gtg agg ctc tgt aaa cca tgc aag ctt gag cca gag ccc cgc ctt 4129
394 tgg gtg gtg cct ggg gca ctc cca cag gtg 4159
397 <210> SEQ ID NO: 9
398 <211> LENGTH: 23

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/599,087

DATE: 07/05/2000

TIME: 18:53:56

Input Set : A:\00450seq.txt

Output Set: N:\CRF3\07052000\I599087.raw

L:11 M:270 C: Current Application Number differs, Replaced Application Number
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:260 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:8
L:260 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:8
L:384 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:8
M:340 Repeated in SeqNo=8